A C T A   U N I V E R S I T A T I S   L O D Z I E N S I S
FOLIA OECONOMICA 123, 1992

*Ene-Margit Tiit**

ESTIMATION OF FUNCTIONS MOMENTS

1. INTRODUCTION

Most statistics, used in multivariate statistical analysis, are functions of moment matrices (including covariance and correlation matrices). In applied statistics, one of the most essential problems is to estimate these statistics on the basis of finite (not very large) sample.

Practically, in many cases, the corresponding functions of empirical moments are used as estimates of functions of theoretical moments. It is well known that these estimates are asymptotically unbiased, but the problem of estimating the bias in the case of finite sample is, in general, not solved.

The aim of the report is

1) to find the expression

$$Et(\mathcal{X}) = \sum_{i=0}^{\infty} A_i n^{-i} \qquad (1)$$

where $t(\mathcal{X})$ is a matrix-formed statistics where all the elements are some moments' functions, A_i are fixed matrices of theoretical moments; here, the term A_i defines the bias of order i, i = = 1, 2, ...; A_0 is the leading term of the estimation (asymptotically unbiased estimator).

2) to construct the family of estimators $B_i(\mathcal{X})$ for a given matrix function τ of theoretical moments (correlation or covariance matrices), fulfilling the conditions

* Professor at the University of Tartu, Tartu.

$$\tau - EB_i(\mathfrak{X}) = 0(n^{-i}), \quad i = 1, 2, \ldots, \tag{2}$$

then $B_i(\mathfrak{X})$ is a n^{-i}-biased estimator for τ.

We shall consider the special case when t and τ are entire rational functions of the sample (correspondingly theoretical) moments of the arbitrary finite order. From here it follows that all results are useable for functions of variance covariance matrix, but to use them for correlation matrix the Taylor's expansion of the function must be exploited.

There exists a lot of solutions of problems noticed for special functions of moments, but in most cases there are some important restrictions (see Herzel, 1987, Cramer, 1946) for instance:

1) the order of the highest moment is restricted;

2) the type of initial distribution is assumed to be known (initial distribution $N(m, \sigma)$);

3) instead of matrices, the univariate statistics are regarded

2. ESTIMATION OF UNIVARIATE MOMENTS' PRODUCT

Let X be given variable, having p-order distribution $(p \in N)$. Let v be a given partition of integer p (see Andrews 1976) $v = (v_1, \ldots, v_s)$, $v_i > 0$, $\Sigma v_i = p$; s denotes the number of parts of the partition v, $s(v) = s$. Let $\mathcal{V}$ be a set of all possible partitions of integer p.

Every partition v defines a product of moments for variable X:

$$\mu(v) = \prod^{s} EX^{v_i} = \prod_{i=1}^{s} \mu_{v_i} \tag{3}$$

as well as for fixed sample $\mathfrak{X}$ of size n the product of empirical moments

$$m(v) = \prod_{i=1}^{s} \left(\frac{1}{n} \sum_{j_i=1}^{n} x_{j_i}^{v_i}\right) = \prod_{i=1}^{s} m_{v_i} \tag{4}$$

For product m(v) the expression (1) has the form

$$Em(v) = \sum_{j=0}^{s-1} n^{-j} \sum_{u=0}^{j} f_{j-u}^{s-u} \sum_{v' \in \mathcal{V}} B(v, v')\, \mu(v') \tag{5}$$

where $B(v, v')$ is the number of different possibilities of getting partition v' by adding the parts of partition v. Note that $\{f_i^l\}$ are the coefficients of the expression

$$F(l) = \sum_{i=0}^{l-1} f_i^l n^{-i}$$

where

$$F(l) = \begin{cases} 1, & \text{if } l = 1 \\ (1 - 1/n) \ldots (1 - (l - 1)/n), & \text{if } l > 1 \end{cases} \qquad (6)$$

All the constants by $\mu(v')$ are standard, independent of the initial distribution of X.

By (5) and

$$Em(v) = \mu(v) + o(n^{-1}), \qquad (7)$$

the step by step procedure for calculation of estimator for the product $\mu(v)$ with bias of given order $B_k(\mathfrak{X})$, see (2), is described in (T i i t 1988, T i i t 1986). Since all estimators $B_k(\mathfrak{X})$ have the form

$$B_k(\mathfrak{X}) = \tilde{\sum_{i=0}} n^{-i} \; C_{k_i} \sum_{v \in \mathcal{V}} m_v$$

it follows from (5) that their expectations have the form

$$\tilde{\sum_{i=0}} n^{-i} \, G_{k_i} \sum_{v \in \mathcal{V}} \mu_v,$$

and from the convergence of the series

$$\tilde{\sum_{k=1}} E(B_k(\mathfrak{X}) - B_{k-1}(\mathfrak{X})) = \tilde{\sum_{i=0}} n^{-i} \tilde{\sum_{k=1}} (G_{k_i} - G_{k_{i-1}}) \sum_{v \in \mathcal{V}} \mu_v \qquad (8)$$

the existance of unbiased estimator $B_-(\mathfrak{X})$ for product $\mu(v)$ follows. The necessary condition of existance of $B_-(\mathfrak{X})$.

$$n > C(p, s)$$

where p is the sum of degrees of product $\mu(v)$ and s the number of factors in it. The existance of the estimator $B_-(\mathfrak{X})$ does not depend on the initial distribution of $\mathfrak{X}$.

It is important for practical applications that the variance of estimators $B_1(\mathfrak{X})$, ..., and $B_-(\mathfrak{X})$ all have the same mean term (of order n^{-1}).

3. ESTIMATION OF MIXED MOMENTS' PRODUCT

Let us regard the h-variate random vector X, $X = (X_1, \ldots, X_h)$ and let

$$\mu(\Lambda) = \mu_{l_1^1 \ldots l_h^1} \mu_{l_1^2 \ldots l_h^2} \ldots \mu_{l_1^g \ldots l_h^g}$$

be the product of g mixed moments $\mu_{l_1^j \ldots l_h^j}$ $(j = 1, \ldots, g)$,

$$\mu_{l_1^j \ldots l_h^j} = EX_1^{l_1^j} \ldots X_h^{l_h^j}$$

To describe product $\mu(\Lambda)$ we must use the bivariate partition of number p $(p = \sum_{i=1}^{h} \sum_{j=1}^{g} l_i^j,\ l_i^j \geq 0)$, having the form of the following matrix:

$$\Lambda = \begin{pmatrix} l_1^1 & l_1^2 & l_1^g \\ l_2^1 & l_2^2 & l_2^g \\ l_h^1 & l_h^2 & l_h^g \end{pmatrix}$$

Let us denote $\sum_{j=1}^{g} l_i^j = l_i^o$, $i = 1, \ldots, h$ and say, vector $l = (l_1^o, \ldots, l_h^o)$ is the marginal vector of partition Λ. Let $\mathfrak{L}$ be the set of all partitions of number p having the marginal vector equal to l.

Then we are able to use the formula (5) for calculating the expectation $Em(\Lambda)$ in the following form

$$Em(\Lambda) = \sum_{j=0}^{s-1} n^{-j} \sum_{u=0}^{j} f_{j-u}^{s-u} \sum_{\Lambda' \in \mathfrak{L}} B(\Lambda, \Lambda')\mu(\Lambda'),$$

where $B(\Lambda, \Lambda')$ is the number of possibilities of getting Λ from Λ' by adding parts (columns) of Λ.
From here, analogously, the estimations of given order can be derived.

4. ESTIMATION OF MATRICES OF MIXED MOMENT'S PRODUCTS

Let X be a random vector $X = (X, \dots, X_m)$ and let $\chi(\Lambda)$ be $q \times r$ -matrix of products of mixed moments of X, having the following elements:

$$\chi_{ij} = \mu(\Lambda, X_{I_{ij}})$$

where I_{ij} is a subvector of h size (fixed by Λ) of the random vector X.

Then all elements of the matrix χ are products of mixed moments of vector X, consisting, in general, of different components, but all having the same form.

Analogously, let $H(\Lambda)$ be the matrix of products of empirical moments of X, having elements

$$h_{ij} = m(\Lambda, X_{ij}), \quad i = 1, \dots, q; \quad j = 1, \dots, r.$$

Then for the expectation of matrix H the formula (9) is applicable, where instead of moments $\mu(\Lambda')$, the matrices $\chi(\Lambda')$ are replaced.

Notice that most of matrix formed statistics used in multivariate statistical analysis have the form of matrix $H(\Lambda)$ or are expressed by the sum (linear combination) of such matrices. From here it follows that for all these matrix-formed statistics the estimators of given order k or unbiased estimators (in the case of sample size, fulfilling the condition (9)), can be constructed.

REFERENCES

A n d r e w s G. (1976), *Partition Theory*, London.

H e r z e l A. (1987), *Multivariate moments and products of moments derivation of expectations and estimators from univariate formulae in simple random sampling*, "Statistica", Vol. 1.

C r a m e r H. (1946), *Mathematical methods in statistics*, Princeton.

T i i t E. M. (1988); *Unbiased and n^{-k}-biased estimations of entire rational functions of moments*, "Tartu Riikliku Ulikooli Toimetised", No. 798, p. 3-18.

T i i t E. M. (1986), *Wychislenią matematicheskikh ożidanij niekotorykh funkcij empiricheskikh momentov*, "Trudy WC TGUn", N° 53, p. 60-85.

Ene Margit Tiit

ESTYMACJA FUNKCJI MOMENTÓW

Problem obciążenia funkcji momentów pojawia się wówczas, gdy próba jest skończona i nie jest duża (np. 10, ..., 100), a szczególnie wtedy, gdy są stosowane momenty wyższych rzędów.

Wyrażenie definiujące obciążenie funkcji momentów ma następującą postać:

$$\sum_{i=1}^{n} n^{-i} A_i,$$

gdzie n jest licznością próby, a współczynnik A_i zależy od momentów teoretycznych i jest dany dla dowolnej wymiernej funkcji (o wartościach całkowitych) momentów empirycznych. Jeżeli τ jest wymierną funkcją momentów teoretycznych, to nieobciążony estymator tej funkcji przybiera postać następującego skończonego iloczynu

$$t = \prod \frac{f_i m(r_i)}{n - 1}$$

gdzie f_i są współczynnikami niezależnymi od początkowego rozkładu, a $m(r_i)$ są iloczynami danych momentów empirycznych. Istnienie estymatora t zależy jedynie od liczności próby n.

Stosując rozwinięcie Taylora, możemy skonstruować dla szerokiej klasy momentów estymatory o danym rzędzie obciążenia;

$$t = 0\ (n^{-h}),$$

gdzie $n \in N$ i jest skończoną liczbą.

Wszystkie otrzymane wyniki są również prawdziwe w przypadku wielowymiarowym.